

HEAD AREA EXTRACTION DEVICE AND REAL-TIME EXPRESSION TRACING DEVICE

Publication number: JP2003108980

Publication date: 2003-04-11

Inventor: TANAKA SHOJI; TANAKA SATOSHI

Applicant: MITSUBISHI ELECTRIC CORP

Classification:

- international: G06T1/00; G06T7/00; G06T7/20; H04N7/14; H04N7/14;
G06T1/00; G06T7/00; G06T7/20; H04N7/14; H04N7/14;
(IPC1-7): H04N7/14; G06T1/00; G06T7/00; G06T7/20

- european:

Application number: JP20010304116 20010928

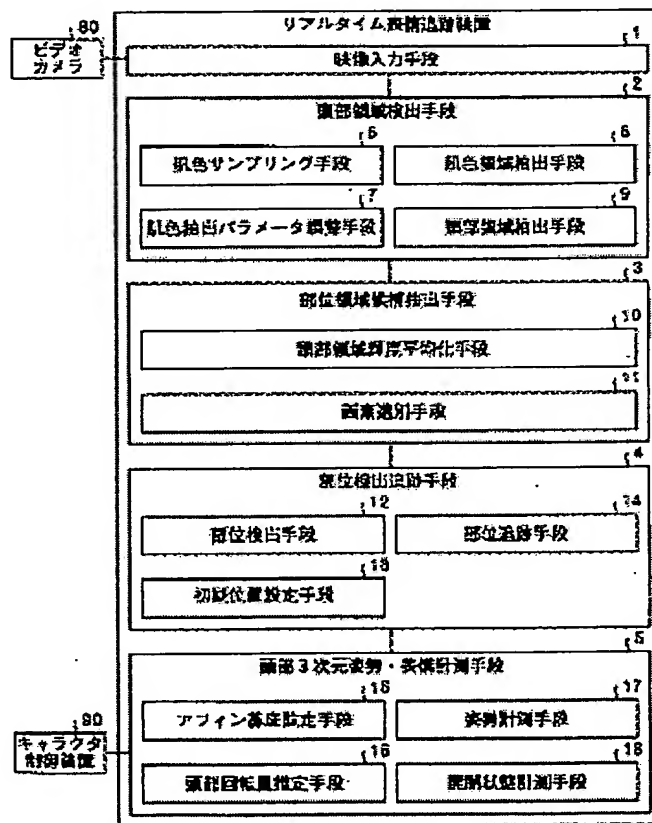
Priority number(s): JP20010304116 20010928

Report a data error here

Abstract of JP2003108980

PROBLEM TO BE SOLVED: To measure the three-dimensional movement of the head in real time from the face image of an unspecified person photographed under an arbitral illumination condition, even by a hardware having low calculation performance.

SOLUTION: This device has a flesh color sampling means 6 sampling the image of a part of the face of a target person under present illumination environment, a flesh color extraction parameter adjustment means 7 adjusting a flesh color extraction parameter for extracting flesh color on the basis of sampling data of the sampling means 6, a flesh color area extraction means 8 extracting a flesh color area from a capture image by use of the flesh color extraction parameter adjusted by the parameter adjustment means 7, and a head area extraction means 9 extracting a head area by extracting a maximum area from the flesh color area extraction result.



Data supplied from the esp@cenet database - Worldwide

①/2

p2893

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-108980

(P2003-108980A)

(43) 公開日 平成15年4月11日 (2003.4.11)

(51) Int.Cl. ⁷	識別記号	F I	テーマコード* (参考)
G 0 6 T 1/00	3 4 0	G 0 6 T 1/00	3 4 0 A 5 B 0 5 7
7/00	1 0 0	7/00	1 0 0 C 5 C 0 6 4
7/20	3 0 0	7/20	3 0 0 B 5 L 0 9 6
// H 0 4 N 7/14		H 0 4 N 7/14	

審査請求 有 請求項の数10 O L (全 26 頁)

(21) 出願番号 特願2001-304116(P2001-304116)

(22) 出願日 平成13年9月28日 (2001.9.28)

(71) 出願人 000006013

三菱電機株式会社

東京都千代田区丸の内二丁目2番3号

(72) 発明者 田中 昭二

東京都千代田区丸の内二丁目2番3号 三

菱電機株式会社内

(72) 発明者 田中 聡

東京都千代田区丸の内二丁目2番3号 三

菱電機株式会社内

(74) 代理人 100089118

弁理士 酒井 宏明

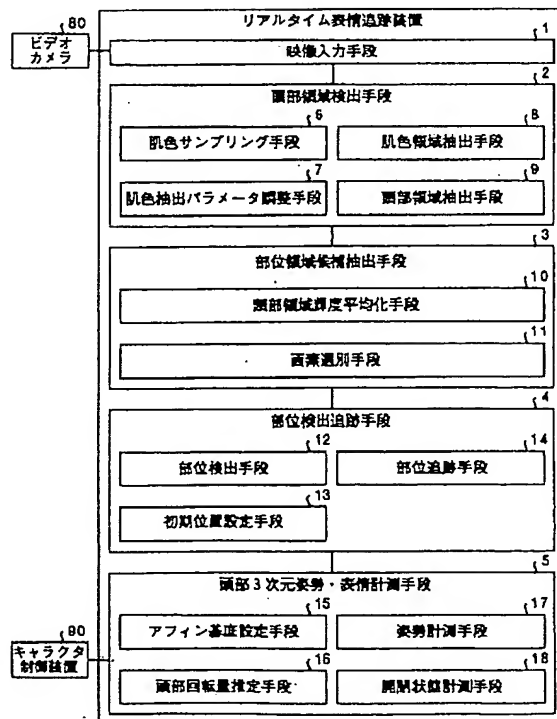
最終頁に続く

(54) 【発明の名称】 頭部領域抽出装置およびリアルタイム表情追跡装置

(57) 【要約】

【課題】 任意の照明条件で撮影された、不特定人物の顔画像から、計算能力が低いハードウェアでも実時間で頭部の3次元的な動きを計測し得ること。

【解決手段】 現照明環境下での対象人物の顔の一部の画像をサンプリングする肌色サンプリング手段6と、この肌色サンプリング手段6のサンプリングデータに基づき、肌色抽出のための肌色抽出パラメータを調整する肌色抽出パラメータ調整手段7と、この肌色抽出パラメータ調整手段7によって調整された肌色抽出パラメータを用いてキャプチャ映像から肌色領域を抽出する肌色領域抽出手段8と、肌色領域抽出結果から最大領域を抽出することにより頭部領域を抽出する頭部領域抽出手段9とを備える。



OP: 現フレームにおける候補領域のみを非肌色に対応する論理値レベルとした部位領域マスク画像と、前フレームにおける部位領域のみを非肌色に対応する論理値レベルとした部位領域マスク画像との排他的論理和を求めたときに、画素値が非肌色に対応する論理値レベルとなる画素数、

D: 前フレームにおける部位領域の中心と候補領域の中心との距離、評価値Eが最も小さな候補領域を各部位領域として特定することを特徴とする請求項7に記載のリアルタイム表情追跡装置。

【請求項9】 順次所定のフレームレートで入力される映像をキャプチャする映像入力手段と、
前記キャプチャした画像から頭部画像を抽出する頭部領域検出手段と、

前記抽出した頭部領域から両目および口を含む各部位の候補領域を抽出する部位領域候補抽出手段と、
抽出した候補領域の中から各部位の位置を検出する部位検出追跡手段と、

前記検出した両目、口の検出位置に基づいて頭部の3次元姿勢を計測するとともに、両目および口の開閉状態を計測する頭部3次元姿勢・表情計測手段とを備え、前記計測した頭部の3次元姿勢および両目および口の開閉状態に基づいてCGキャラクタの動きを制御するリアルタイム表情追跡装置であって、

前記頭部3次元姿勢・表情計測手段は、
最初に検出した両目および口の位置から3次元空間上の仮想平面を設定するアフィン基底設定手段と、
前記検出した両目および口位置から頭部の左右および上下方向の回転量を推定する頭部回転量推定手段と、
前記検出した両目および口位置から得た4点の座標を結ぶ矩形を前記推定した頭部の左右および上下方向の回転量を用いて歪ませ、該歪ませた矩形の4点の座標を用いて頭部の3次元姿勢を推測する姿勢計測手段と、
頭部の動きに応じて両目および口の開閉状態を推測する開閉状態計測手段と、

を備えることを特徴とするリアルタイム表情追跡装置。

【請求項10】 前記頭部回転量推定手段は、両目領域を結ぶ直線をX軸とし、X軸に垂直で口領域の中心位置を通る直線をY軸とした頭部のローカル座標系を設定し、このローカル座標系において求めた頭部領域の外接矩形の左右の辺と片目との相対距離と、外接矩形の上下の辺と口領域との相対距離から頭部の左右、上下方向の回転量をそれぞれ推定することを特徴とする請求項9に記載のリアルタイム表情追跡装置。

【請求項11】 前記開閉状態計測手段は、前記姿勢計測手段によって推定した頭部の3次元姿勢情報を用いて対象人物が正面を向いたときの両目および口領域を再現することに基づき両目および口の開閉状態を計測することを特徴とする請求項9または10に記載のリアルタイム表情追跡装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、本人の顔を送信する代わりにCGキャラクタの映像を相手に送信することによって人物映像を互いに通信するテレビ電話など通信システムに適用され、特にカメラによって撮像された顔の映像から頭部の3次元的な姿勢情報と顔の表情を計測し、この計測結果に基づいてCGキャラクタの動きを制御する代理応答によるリアルタイム表情追跡装置に関するものである。

【0002】

【従来の技術】 例えば、図30は、特開2000-331190号公報に示された従来の仮想変身装置（第1の従来技術）を示すものであり、この仮想変身装置は、顔画像を入力するビデオカメラと、ビデオカメラを回転させる電動雲台と、ビデオカメラから入力された顔画像から顔の軸の回転、あるいは顔の軸周りの回転と視線方向を検出し、両目および口の形状変化を検出する顔画像認識装置と、この計測結果に基づいてCG（コンピュータグラフィックス）で構築された仮想空間のキャラクタを制御する仮想環境合成装置とを備えている。

【0003】 この第1の従来技術では、ビデオカメラから入力された顔画像を、予め設定したRGB空間上に構築された肌色モデルに従って肌色を1、肌色以外を0とする2値化処理を行う。次に、2値化した顔領域の重心を求め、重心が画像の中心になるように電動雲台装置を制御し、カメラのアングルを修正する。次に、重心位置に基づき顔領域内に存在する穴を両目および口として検出する。次に、予め設定したテンプレートをを用いたテンプレートマッチングにより目領域を追跡し、黒目の位置から視線方向を求める。また、両目を結んだ直線と画像の水平軸との角度を計測し、さらに、両目間の距離から、顔の軸周りの回転を検出する。そして、両目および口の周囲画像を離散コサイン変換したときの各周波数帯域での電力変化を捉えることで、両目および口の形状変化を計測する。以上の計測結果に基づいてCGで構築された仮想空間のキャラクタの頭部および表情を制御する。

【0004】 また、特開2000-259831号公報の表情検出装置（第2の従来技術）では、連続する各フレームの画像において、選択した複数の特徴点を追跡し、各フレーム毎に前記複数の特徴点を頂点とするドロネー網を構成し、このドロネー網を用いて表情筋モデルを特徴点の移動に基づき変位させることにより、表情筋モデルの変化を求めるようにしている。

【0005】 また、特開平11-306348号公報（第3の従来技術）においては、大きさが固定のウィンドウマスクを画像全体に走査し、マスク内の輝度分散を正規化することにより、照明条件が変化しても安定して対象物の特徴量を抽出可能とした対象物検出装置に関する

10

20

30

40

50

域中の両目および口の候補領域を抽出する画素選別手段とを備えることを特徴としている。

【0017】この発明によれば、頭部領域検出手段によって抽出された頭部領域の輝度を平均化・正規化し、該輝度平均化・正規化された画像を用いて頭部領域中の両目および口の候補領域を抽出するようにしている。

【0018】つぎの発明にかかるリアルタイム表情追跡装置は、上記の発明において、前記頭部領域輝度平均化手段は、頭部領域を複数の小領域に分割し、各小領域毎にヒストグラム平均化処理を行うことを特徴とする。

【0019】この発明によれば、頭部領域を複数の小領域に分割し、各小領域毎にヒストグラム平均化処理を行うようにしている。

【0020】つぎの発明にかかるリアルタイム表情追跡装置は、上記の発明において、前記ヒストグラム平均化処理では、所定の閾値を越えた頻度をもつ画素値の頻度を他の画素値に分散させる処理を加えることを特徴とする。

【0021】この発明によれば、ヒストグラム平均化処理では、所定の閾値を越えた頻度をもつ画素値の頻度を他の画素値に分散させるようにしている。

【0022】つぎの発明にかかるリアルタイム表情追跡装置は、順次所定のフレームレートで入力される映像をキャプチャする映像入力手段と、前記キャプチャした画像から頭部画像を抽出する頭部領域検出手段と、前記抽出した頭部領域から両目および口を含む各部位の候補領域を抽出する部位領域候補抽出手段と、抽出した候補領域の中から各部位の位置を検出する部位検出追跡手段と、前記検出した両目、口の検出位置に基づいて頭部の3次元姿勢を計測するとともに、両目および口の開閉状態を計測する頭部3次元姿勢・表情計測手段とを備え、前記計測した頭部の3次元姿勢および両目および口の開閉状態に基づいてCGキャラクタの動きを制御するリアルタイム表情追跡装置であって、前記部位検出追跡手段は、前記部位領域候補抽出手段によって検出された両目、口の候補領域から両目および口の領域を夫々特定する部位検出手段と、前記部位検出手段によって、両目および口の領域が特定できない場合に、現フレームで特定された部位領域の位置と、この特定された部位領域の前フレームでの位置とを用いて移動ベクトルを求め、この移動ベクトルを用いて前記特定できなかった部位の位置を特定する部位追跡手段とを備えることを特徴とする。

【0023】この発明によれば、部位検出手段によって両目および口の領域が特定できない場合には、現フレームで特定された部位領域の位置と、この特定された部位領域の前フレームでの位置とを用いて移動ベクトルを求め、この移動ベクトルを用いて特定できなかった部位の位置を特定している。

【0024】つぎの発明にかかるリアルタイム表情追跡装置は、上記の発明において、前記部位検出手段は、前

フレームについての部位領域の中心座標を中心に一定の大きさの矩形領域を設定し、その矩形領域中に存在する現フレーム候補領域を求め、求めた候補領域夫々について、判別式 $E = |SP - SC| + OP + D$ を用いて評価値 E を夫々取得し、

SP：前フレームにおける部位領域の画素数、

SC：現フレームにおける候補領域の画素数、

OP：現フレームにおける候補領域のみを非肌色に対応する論理値レベルとした部位領域マスク画像と、前フレームにおける部位領域のみを非肌色に対応する論理値レベルとした部位領域マスク画像との排他的論理和を求めたときに、画素値が非肌色に対応する論理値レベルとなる画素数、

D：前フレームにおける部位領域の中心と候補領域の中心との距離、

評価値 E が最も小さな候補領域を各部位領域として特定することを特徴としている。

【0025】この発明によれば、前フレームについての部位領域の中心座標を中心に一定の大きさの矩形領域を設定し、その矩形領域中に存在する現フレーム候補領域を求め、求めた候補領域夫々について、判別式 $E = |SP - SC| + OP + D$ を用いて評価値 E を夫々取得し、評価値 E が最も小さな候補領域を各部位領域として特定する。

【0026】つぎの発明にかかるリアルタイム表情追跡装置は、順次所定のフレームレートで入力される映像をキャプチャする映像入力手段と、前記キャプチャした画像から頭部画像を抽出する頭部領域検出手段と、前記抽出した頭部領域から両目および口を含む各部位の候補領域を抽出する部位領域候補抽出手段と、抽出した候補領域の中から各部位の位置を検出する部位検出追跡手段と、前記検出した両目、口の検出位置に基づいて頭部の3次元姿勢を計測するとともに、両目および口の開閉状態を計測する頭部3次元姿勢・表情計測手段とを備え、前記計測した頭部の3次元姿勢および両目および口の開閉状態に基づいてCGキャラクタの動きを制御するリアルタイム表情追跡装置であって、前記頭部3次元姿勢・表情計測手段は、最初に検出した両目および口の位置から3次元空間上の仮想平面を設定するアフィン基底設定手段と、前記検出した両目および口位置から頭部の左右および上下方向の回転量を推定する頭部回転量推定手段と、前記検出した両目および口位置から得た4点の座標を結ぶ矩形を前記推定した頭部の左右および上下方向の回転量を用いて歪ませ、該歪ませた矩形の4点の座標を用いて頭部の3次元姿勢を推測する姿勢計測手段と、頭部の動きに応じて両目および口の開閉状態を推測する開閉状態計測手段とを備えることを特徴とする。

【0027】この発明によれば、最初に検出した両目および口の位置から3次元空間上の仮想平面を設定し、検出した両目および口位置から頭部の左右および上下方向

ためのフローチャートである。図3は、図1のリアルタイム表情追跡装置のトラッキングフェーズの動作の概要を説明するためのフローチャートである。これら図2および図3を用いてリアルタイム表情追跡装置の動作の概略を説明する。

【0041】リアルタイム表情追跡装置で行われる動作手順には、頭部の動きを追跡するための情報として両目および口の位置および無表情時の状態等を取得するキャリブレーションフェーズと、実際に頭部の動きおよび両目および口を追跡し、頭部姿勢と両目および口の開閉状態つまり表情を計測するトラッキングフェーズがある。

【0042】キャリブレーションフェーズでは、まず、映像入力手段1によってビデオカメラ80からの映像をキャプチャする(ステップS100)。なお、人物の映像をビデオカメラ80で撮像する際に、ユーザに対して「カメラに対して正面を向き、両目を開け、口を閉じる」ように指示することで、無表情時の人物映像を得る。つぎに、頭部領域検出手段2において、撮影環境下におけるユーザの肌色をサンプリングし(ステップS110)、このサンプリングデータを用いて予め設定した肌色抽出パラメータの調整を行う(ステップS120)。そして、調整した肌色抽出パラメータを用いて実際に肌色領域を抽出し(ステップS130)、抽出した領域の中から頭部領域を検出する(ステップS140)。次に、部位領域候補抽出手段3において、抽出した頭部領域から両目、口の候補領域を抽出し(ステップS150)、部位検出追跡手段4において両目領域および口領域をそれぞれ検出し(ステップS160)、検出した両目および口領域から各部位の位置、大きさ、テンプレートの初期値を記憶する(ステップS170)。最後に、頭部3次元姿勢・表情計測手段5において、求めた両目および口の位置に基づき、トラッキングフェーズにおいて頭部の3次元姿勢情報を求めるためのアフィン基底(3次元空間上の仮想点)を設定する(ステップS180)。

【0043】トラッキングフェーズでは、映像入力手段1によってビデオカメラ80からの映像をキャプチャする(ステップS200)。頭部領域検出手段2においては、キャリブレーションフェーズで設定した肌色抽出パラメータを用いてキャプチャした映像中から肌色を抽出し(ステップS210)、抽出した領域から頭部領域を検出する(ステップS220)。次に、部位領域候補抽出手段3において、両目および口の候補領域を抽出する(ステップS230)。つぎに、部位検出追跡手段4は、前フレームで検出した両目および口位置に基づき、部位領域候補抽出手段3で抽出した候補領域の中から現フレームにおける両目および口領域を検出する(ステップS240)。次に、頭部3次元姿勢・表情計測手段5において、部位検出追跡手段4で検出した両目および口位置(2次元画像点)と予め設定した3次元空間上の仮

想点から頭部の3次元姿勢情報を計測し(ステップS250)、その計測情報に基づいて両目および口の開閉状態を計測する(ステップS260)。最後に、計測した両目および口の開閉状態情報及び頭部の姿勢情報はキャラクタ制御装置90に入力され、キャラクタ制御装置90によってCGキャラクタの頭部の動きおよび表情が制御される(ステップS270)。

【0044】[キャリブレーションフェーズ]次に、図1のリアルタイム表情追跡装置のキャリブレーションフェーズにおける動作を図4～図17を用いて詳細に説明する。

【0045】(a)頭部領域検出手段2での処理
まず、図4～図10を用いて頭部領域検出手段2が行う図2のステップS110～S140の処理の詳細について説明する。

【0046】図4は、頭部領域検出手段2における肌色サンプリング手段6の動作を説明するための図である。図5は、肌色サンプリング手段6および肌色抽出パラメータ調整手段7の動作を説明するためのフローチャートである。

【0047】まず、使用する照明環境下におけるユーザの肌色をサンプリングするために、図4に示すように、キャプチャ映像19に重ねて、サンプリング領域を指定するためのサンプリングウィンドウ20を表示する(ステップS300)。次に、ユーザは、マウスあるいはその他のポインティングデバイスやキーボード等を用いて、サンプリングウィンドウ20を頬あるいは頬などの肌色のみ抽出可能な位置に移動させ、サンプリング可能であることをシステムに伝える(ステップS310)。なお、最初に表示したサンプリングウィンドウ20の位置に合わせてユーザ自身が頭を動かして位置を調整しても良い。

【0048】次に、サンプリングウィンドウ20内の全ての画素の色を肌色抽出のための色空間(肌色モデル空間)に写像し(ステップS320)、写像画素の写像空間での最大値および最小値を用いて予め設定した肌色抽出パラメータを調整する(ステップS330)。

【0049】ここで、肌色抽出空間は、例えば、輝度変化に比較的口バストな色空間を新たに構築するとか、画素の色データ空間(R、G、B空間)上で構築するなどの方法を用いる。ここでは、下記のような、輝度変化に比較的口バストな色空間を用いることにする。

【0050】R(レッド)、G(グリーン)、B(ブルー)を各画素の色の3原色の成分だとすると、まず、次式により色を正規化する。

【0051】

$$c1 = \arctan(R / \max(G, B)) \cdots \cdots \text{式(1)}$$

$$c2 = \arctan(G / \max(R, B)) \cdots \cdots \text{式(2)}$$

$$c3 = \arctan(B / \max(R, G)) \cdots \cdots \text{式(3)}$$

【0052】上記式で正規化した色をさらに次式で変換

域のコントラストを一定に保つ処理を行う。まず、頭部領域輝度平均化手段10は、頭部領域の外接矩形を求め、その外接矩形領域を例えば8×8の小領域に分割する(ステップS500)。つぎに、頭部領域輝度平均化手段10は、各小領域毎にヒストグラム平均化処理を行う(ステップS510)。

【0067】ヒストグラム平均化処理は、次のようにして行う。まず、各小領域毎に画素値と頻度の関係を示すヒストグラムを求める。次に、累積頻度(頻度の各階級(画素値)までの累計)を求め、各累積頻度を累積頻度の最大値で割って、各累積頻度の比率を求める。そして、求めた比率に小領域内の画素値の最大値を掛け合わせ、四捨五入により小数点以下を丸める。ここで得られた値が、平均化後の画素値となる。最後に、平均化後の画素値の頻度を、平均化前の頻度から求める。

【0068】例えば、図13に示すように小領域内の画素値が0から7の範囲内にあり、その頻度が図13に示す通りであった場合、平均化後のそれぞれの画素値の頻度は図14に示す通りになる。例えば、平均化後の画素値が4の場合、画素値4に対応する平均前の画素値は2と3であるため、その頻度は、9+2=11となる。

【0069】ここで、上記のとおり適応型ヒストグラム平均化法では、特にコントラストが低い小領域において、領域内の大半の画素値がヒストグラムの極大点に割り当てられることから、ノイズが多く発生する可能性がある。そこで、図12(a)に示すようにある閾値を超えた頻度をもつ画素値31が存在する場合には、図12(b)に示すように、それらの頻度を他の画素値に分散させる処理を行うようにしており、これによりノイズの発生を抑えることが可能である。

【0070】以上の処理により、常に一定のコントラストを得られることから、画素選別手段11では、一定の閾値thaを用い、頭部領域内の輝度値が閾値tha以下の画素(暗い画素)を論理レベル1とし、それ以外を*

$$D_{i,j} = \begin{cases} \infty : f_{i,j} = 1 \\ 0 : f_{i,j} = 0 \end{cases}$$

【0076】Step2. 初期化した画像を左上から右下に※ ※向かって走査し、次の規則で逐次D'_{i,j}を更新する。

$$D''_{i,j} = \min(D'_{i,j}, D''_{i-1,j} + 1, D''_{i,j-1} + 1) \dots \text{式(9)}$$

【0077】Step3. 先のStep2で得られたD''_{i,j}に対して、右下から左上に向かって走査し、次の規則で逐次★

$$D_{i,j} = \min(D'_{i,j}, D''_{i,j} + 1, D''_{i,j+1} + 1) \dots \text{式(10)}$$

【0078】上式(10)によって得られたD_{i,j}が距離画像の各画素データとなる。したがって、これら得られた距離画像から、距離値が最大となる画素を求め、この画素を頭部領域の重心とする。

【0079】距離画像変換の特徴は、領域の形が変化しても安定した重心位置を求めることがある。なお、距離画像変換を用いず、画素の座標値の平均により重心を求めても良い。

* 論理レベル0とし(ステップS520)、さらに、画素値が1の画素を4連結あるいは8連結で結合し領域分割する(ステップS530)。最後に、微小領域を除去することにより、各部位(両目と口と鼻)の候補領域を抽出できる(ステップS540)。

【0071】以上のように、頭部全体を一領域として抽出し、その頭部領域のコントラストを常に一定にする処理を施すことにより、両目や口の部位領域の抽出処理を固定の閾値thaを用いて実行することができる。したがって、高速処理が可能となり、かつ輝度変化に頑強なシステムを構築することができる。

【0072】(c)部位検出追跡手段4での処理次に、図15および図16を用いて部位検出追跡手段4がキャリブレーションフェーズにおいて行う図2のステップS160およびS170の動作を説明する。図15は、キャリブレーションフェーズにおける部位検出追跡手段4の動作を説明するためのフローチャートである。

【0073】まず、部位検出手段12は、頭部領域検出手段2で抽出した頭部領域の重心を求める(ステップS600)。この重心位置は、周知の距離変換処理などを用いて求める。

【0074】距離変換処理とは、画像中のオブジェクトの各画素値を、各画素位置から背景領域への最短距離に置き換える変換処理である。距離の概念としては、最も単純な市街地距離(4連結距離)とチェス盤距離(8連結距離)がよく使われる。ここでは、市街地距離を用いたアルゴリズムを説明する。

【0075】Step1. まず、入力画像を二値化した各画素データをf_{i,j}とし、D_{i,j}を初期化変換された多値データとした場合、次のように初期化変換する。すなわち、画素値が1の頭部領域内の画素は、多値データ∞(実際には、100などの大きな値)に置換し、画素値が0の背景画素は、0に置換する。

【数1】

式(8)

$$E = |sp - sc| + op + D$$

【0095】ここで、Eは評価値、SPは前フレームにおける部位領域の画素数、SCは現フレームにおける候補領域の画素数、OPは現フレームにおける候補領域のマスク画像（候補領域の画素のみが1で、それ以外は0の画像）と前フレームにおける部位領域のマスク画像（部位領域の画素のみが1で、それ以外は0の画像）との排他的論理和を求めたときに、画素値が1となる画素数、Dは前フレームにおける部位領域の中心と候補領域の中心との距離である。

【0096】上記式（11）で求めた値Eが最も小さいものを対象領域として選択することにより、前フレームの部位領域の位置を基準とした一定範囲内に存在する現フレームの候補領域の中から対象領域を特定する（ステップS830）。すなわち、図22に示すような小さいノイズ領域47が前フレームの部位領域に完全に包含されたとしても、その場合は式（11）の $|SP - SC|$ とOPの値が大きくなるため、このようなノイズ領域を除去できる。

【0097】このような処理を、左目、右目、口の領域について夫々実行する（ステップS810～S840）。

【0098】以上の処理により全ての部位を検出できた場合は、部位領域マスク画像を、現在のフレームのもので更新し、かつ各部位（左目、右目、口）についての検出領域の中心位置を求め、これを記憶する（ステップS850およびS860）。

【0099】見つからない部位が存在した場合は（ステップS870）、現フレームで検出された部位の移動ベクトルから、検出できなかった部位の現フレームでの位置を予測する。例えば、図23に示すように、現フレームで検出できなかった部位（対象部位）54が存在した場合、現フレームで検出された他の部位48の位置とその部位の前フレームでの位置49からフレーム間の移動ベクトル50を求める。そして、対象部位54の前フレームにおける位置51に、他の部位の検出位置から求めた移動ベクトル50を加算して、現フレームでの推定位置を求める（ステップS890）。そして、求めた位置を含む所定の矩形領域（例えば 16×16 ）53中の画素に着目し、この矩形領域中の画素に対し前述したステップS820およびステップS830の処理を実行することで、対象部位54を検出する（ステップS900）。

【0100】矩形領域53内に全く候補領域が存在しない場合は、顔の傾きなどによる隠れが生じているものとし、ステップS890で推定した位置を現フレームでの対象部位の位置とし、矩形領域53自体をその部位領域として記憶する（ステップS910、S920）。

【0101】なお、ステップS870で、現フレームの

部位領域を全く検出できなかった場合は、部位検出手段12によって図15のステップS600～S640の処理を再度行い、部位領域を再検出する（ステップS880）。

【0102】このように、部位を1つ検出できれば、他の部位を検出漏れしても、検出した部位の移動ベクトルから検出漏れした部位の現フレームでの位置を予測しているので、頑強な部位追跡が行える。さらに、隠れなどにより映像中に対象となる部位が現れない場合でも暫定的な部位領域を設定することから、隠れた部位が出現したときにその部位を即座に追跡可能となり、つまりは、頭部の各部位の滑らかな動きを再現可能となる。

【0103】（d）' 頭部3次元姿勢・表情計測手段5での処理

次に、図24～図29を用いて頭部3次元姿勢・表情計測手段5のトラッキングフェーズにおける動作を詳細に説明する。図24および図27は、トラッキングフェーズにおける頭部3次元姿勢・表情計測手段5の動作を説明するためのフローチャートである。

【0104】まず、頭部回転量推定手段16においては、図25に示すように、部位検出追跡手段4で求められた現フレームの両目領域から、両目の外側の端点70、71を求め、これら端点70、71を結ぶ直線55を求める（ステップS1000）。また、直線55に直交し、口の中心位置59を通る直線56を求める（ステップS1010）。求めた直線55をX軸とし、直線56をY軸とするローカル座標系を設定し、X軸55およびY軸56のそれぞれに平行な辺を持ち、抽出された頭部領域に外接する外接矩形57を求める（ステップS1020）。外接矩形57のX軸方向の辺の長さを1とし、キャリブレーションフェーズで計測した方の目の内側の端点58とY軸に並行な2辺72、73までの相対距離 $L a'$ 、 $L b'$ を夫々求める（ステップS1030）。同様に、外接矩形のY軸方向の長さを1とし、口の中心59とX軸に平行な2辺74、75までの相対距離 $L c'$ 、 $L d'$ を夫々求める（ステップS1040）。

【0105】次に、両目の外側の端点70、71と、端点70、71を通りY軸に平行な直線と口の中心を通りX軸に平行な直線との交点（2点）76、77とでできる矩形60を求める（ステップS1050）。

【0106】ここで、X軸については右方向を正方向、Y軸については上方向を正方向としたとき、片目のX軸正方向の相対距離 $d e c (=L b')$ と、キャリブレーションフェーズで記憶したX軸正方向の相対距離 $d e i (=L b)$ とから次式（12）で頭部の左右方向の回転量を求める。

$$R f i = d e c / d e i \quad \cdots \cdots \text{式（12）}$$

矩形の3次元空間上の上下方向と左右方向の方向ベクトル(X軸、Y軸)を求めることができる。

【0125】平行する辺のカメラ画像平面64上における直線の方程式を

$$a_1 x + b_1 y + c_1 = 0$$

……式(18) *

$$a_1 P_{11} X_c + (a_1 P_{12} + b_1 P_{22}) Y_c$$

$$+ (a_1 P_{13} + b_1 P_{23} + c_1) Z_c = 0 \quad \text{……式(20)}$$

$$a_2 P_{11} X_c + (a_2 P_{12} + b_2 P_{22}) Y_c$$

$$+ (a_2 P_{13} + b_2 P_{23} + c_2) Z_c = 0 \quad \text{……式(21)}$$

【0127】これら2つの平面の法線ベクトル(X、Y、Zの係数)の外積を求めると上記方向ベクトル(X軸、Y軸)を求めることができる。

【0128】以上で、カメラ座標系65における矩形のX軸、Y軸に相当する方向ベクトルを求めることができるが、画像から得られる情報の誤差により、得られた方向ベクトルが図29に示すように直交していない場合がある。そこで、求めた方向ベクトルをS1、S2としたとき、そのベクトルS1、S2を基に直交ベクトルV1、V2を求める。Z軸方向のベクトルは、求めたV1とV2の外積から求められる。この3つの方向ベクトルが式(16)における回転成分Rとなる。

【0129】回転成分Rが分かれば、2次元座標と3次元座標の対応点を式(16)と式(17)に代入することにより並進成分Tを求めることができる。

【0130】姿勢計測手段17では、まず頭部回転量推定手段16で求めた矩形の4頂点の座標から式(18)に示す各辺の直線パラメータ(方程式)を求め(ステップS1100)、求めた直線パラメータを用いて式(20)および式(21)に基づき、アフィン基底設定手段15で設定した仮想3次元平面のX軸、Y軸を求める(ステップS1110)。そして、前述したように、求めた軸が直交するように修正し、更にこの修正したX軸、Y軸からZ軸を求め、これら3軸(X軸、Y軸、Z軸)の方向ベクトルから回転行列(回転成分)Rを求め(ステップS1120)、さらにこの回転成分Rを用いて得られた2次元座標と3次元座標の対応点を式(16)(17)に代入することで、並進行列(並進成分)Tを求める(ステップS1130)。

【0131】以上のようにして求めた投影行列を用いて、実際に3次元空間上の仮想点をカメラ画像平面に投影したときの誤差に応じて投影行列を修正(ステップS1140)、誤差が閾値以下になったときの投影行列を頭部の3次元姿勢情報とし(ステップS1150)、この3次元姿勢情報をキャラクタ制御装置90に出力することで、CGキャラクタの頭部の3次元姿勢を制御する。

【0132】このように、顔画像から検出する両目および口の3点から3次元空間上の矩形(仮想平面)を規定し、追跡時に両目および口の3点から作成した矩形を頭部の動きに応じて歪ませることにより、3次元平面を2

$$* a_2 x + b_2 y + c_2 = 0$$

……式(19)

とすると、カメラ座標系65におけるこれらの各直線を含む3次元の平面の方程式は次式(20)(21)であらわすことができる。

【0126】

10 次元に投影したときの歪みを擬似的に再現し、本来4点以上の3次元と2次元の対応点がなければ求めることができない3次元姿勢情報を画像から得られる両目および口の3点のみで推定するようにしている。

【0133】次に開閉状態計測手段18の動作を説明する。開閉状態計測手段18では、姿勢計測手段17で求めた投影行列、すなわち頭部の3次元姿勢情報を用いて、ユーザが正面を向いたときのカメラ画像における両目および口領域を再現し、再現した領域の画像垂直方向(Y方向)の長さ、初期位置設定手段13に記憶されている初期状態における各部位領域の画像垂直方向の長さとの比率を求める。この比率が、両目および口がどの程度開閉しているかを示す開閉状態情報となる。

【0134】このように3次元姿勢情報を用いてユーザが正面を向いたときのカメラ画像における両目および口領域を推定しているので、例えば頭部が横や上を向いている画像においても正面を向いた場合の画像を推定でき、2次元画像のみから両目および口の開閉状態をより正確に求めることができる。

【0135】このようにして、求められた頭部の3次元姿勢情報および両目および口の開閉状態情報は、キャラクタ制御装置90に入力される。キャラクタ制御装置90は、入力された頭部の3次元姿勢情報および両目および口の開閉状態情報を用いてCGキャラクタの頭部の動きおよび両目および口の開閉状態を可変制御することで、ビデオカメラ80で撮像した利用者の動き、表情に追従させてCGキャラクタの動き、表情をリアルタイムに変化させる。

【0136】

【発明の効果】以上説明したように、この発明によれば、使用する照明環境下で対象人物(利用者)の肌色をサンプリングし、このサンプリングデータを用いて肌色抽出のための肌色抽出パラメータを調整するようにしているので、任意の照明環境あるいは利用者毎の個人差に適應して利用者の頭部領域を確実に抽出することができる。

【0137】つぎの発明によれば肌色領域抽出手段による肌色領域抽出後の2値画像に膨張収縮処理を加えることで、目、鼻、口以外の微小な領域や裂け目が除去された肌色領域を得るようにしているので、頭部領域全体を一領域として抽出することが可能となり、頭部にかかわ

の一例を示した図である。

【図8】 膨張マスクおよび収縮マスクを例示する図である。

【図9】 検出した頭部領域に発生した裂け目を埋める処理を説明するための図である。

【図10】 頭部領域内の全ての穴を埋める論理演算処理を説明するための図である。

【図11】 部位領域候補抽出手段の動作を説明するためのフローチャートである。

【図12】 適応型ヒストグラム平均化法の欠点であるノイズ発生を抑える処理を説明するための図である。

【図13】 適応型ヒストグラム平均化法を説明するための図である。

【図14】 適応型ヒストグラム平均化法を説明するための図である。

【図15】 キャリブレーションフェーズにおける部位検出追跡手段の動作を説明するためのフローチャートである。

【図16】 部位検出手段において両目および口領域を特定する際に用いるマスク領域を示した図である。

【図17】 キャリブレーションフェーズにおける頭部3次元姿勢・表情計測手段5動作を説明するためのフローチャートである。

【図18】 アフィン基底設定手段で設定する3次元空間上の仮想点を示した図である。

【図19】 頭部移動量推定手段で求める両目の端点および口の中心点の頭部領域の外接矩形に対する相対位置を説明するための図である。

【図20】 トラッキングフェーズにおける部位検出追跡手段の動作を説明するためのフローチャートである(その1)。

【図21】 トラッキングフェーズにおける部位検出追跡手段の動作を説明するためのフローチャートである(その2)。

【図22】 部位追跡手段での現フレームにおける部位領域の追跡方法を説明するための図である。

*【図23】 検出できなかった部位領域を検出できた部位領域の位置から予測する処理を説明するための図である。

【図24】 トラッキングフェーズにおける頭部3次元姿勢・表情計測手段の動作を説明するためのフローチャートである。

【図25】 頭部回転量推定手段での左右上下方向の頭部回転量を推定する処理を説明するための図である。

【図26】 頭部回転量推定手段において3次元空間上の仮想点(アフィン基底)に対応する対応点を求める処理を説明するための図である。

【図27】 トラッキングフェーズにおける頭部3次元姿勢・表情計測手段の動作を説明するためのフローチャートである。

【図28】 姿勢計測手段での3次元と2次元の対応点から頭部の3次元姿勢情報を求める処理を説明するための図である。

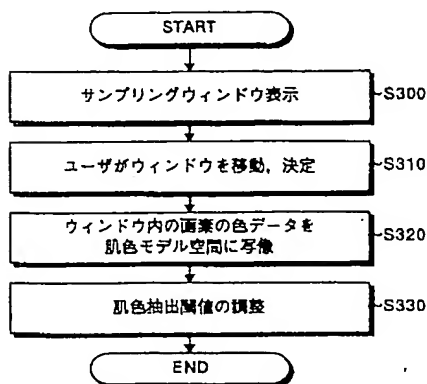
【図29】 姿勢情報を求める際の誤差を補正する処理を説明するための図である。

【図30】 従来技術を示す図である。

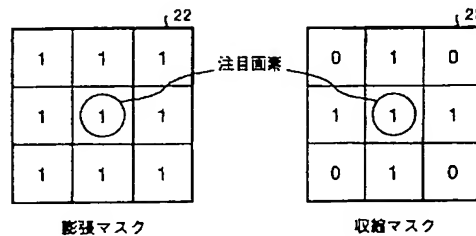
【符号の説明】

1 映像入力手段、2 頭部領域検出手段、3 部位領域候補抽出手段、4 部位検出追跡手段、5 3次元姿勢・表情計測手段、6 肌色サンプリング手段、7 肌色抽出パラメータ調整手段、8 肌色領域抽出手段、9 頭部領域抽出手段、10 頭部領域輝度平均化手段、11 画素選別手段、12 部位検出手段、13 初期位置設定手段、14 部位追跡手段、15 アフィン基底設定手段、16 頭部回転量推定手段、17 姿勢計測手段、18 開閉状態計測手段、20 サンプリングウィンドウ、22 膨張マスク、23 収縮マスク、33 左目マスク、34 右目マスク、35 口マスク、50 移動ベクトル、53 矩形領域、57 外接矩形、64 カメラ画像平面、80 ビデオカメラ、90 キャラクタ制御装置。

【図5】



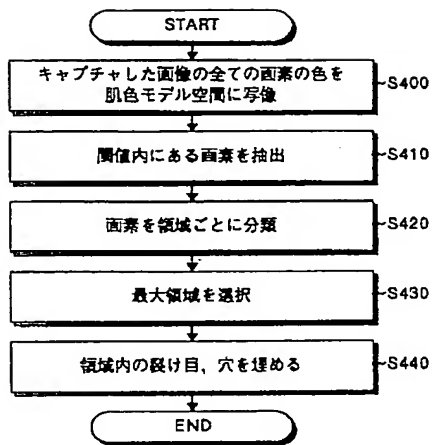
【図8】



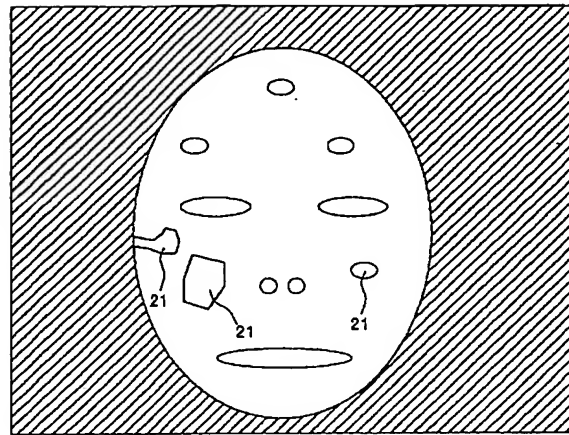
【図14】

画素値	平均化後の画素値
0	0
1	10
2	8
3	0
4	11
5	0
6	15
7	7

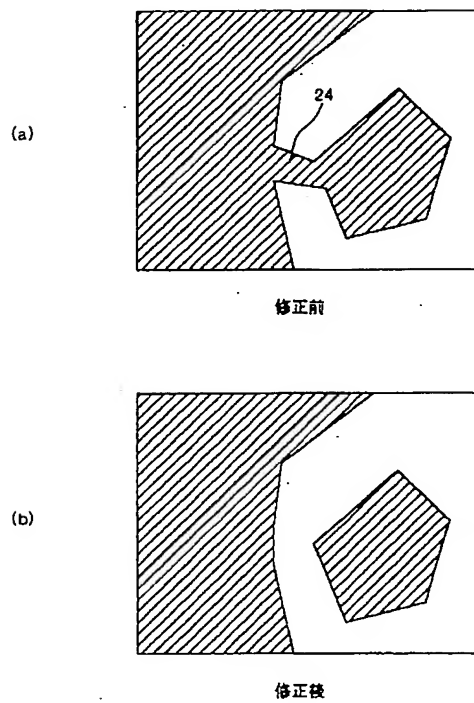
【図6】



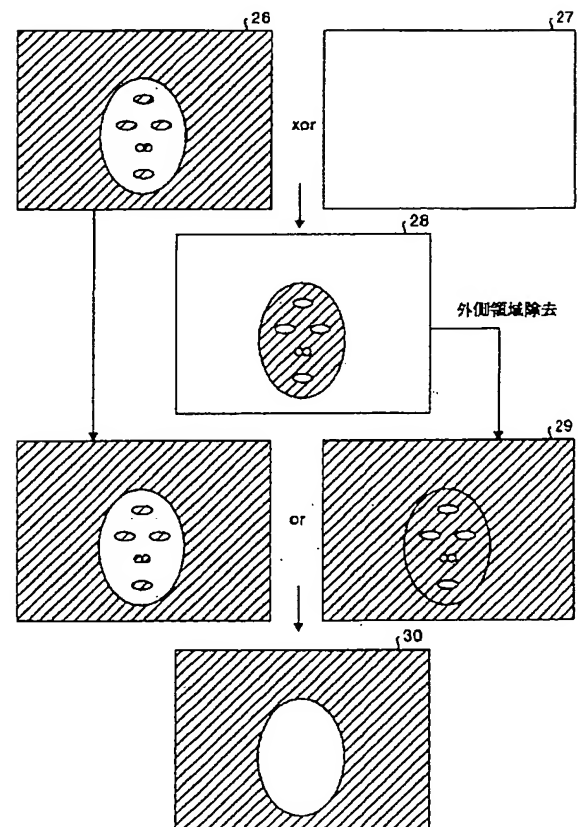
【図7】



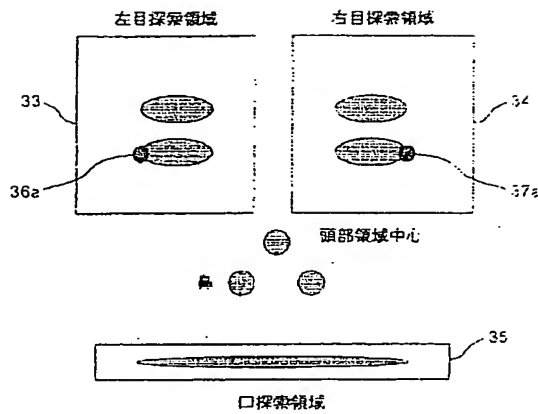
【図9】



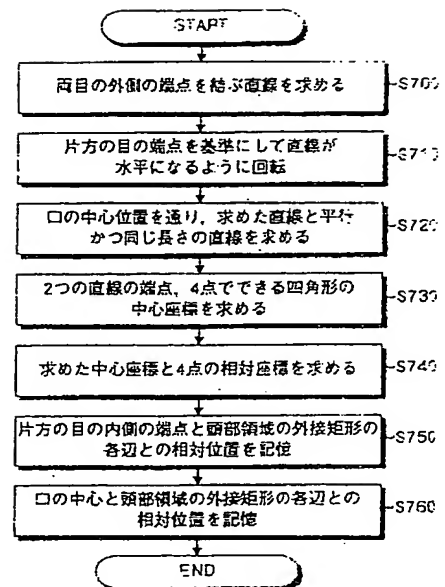
【図10】



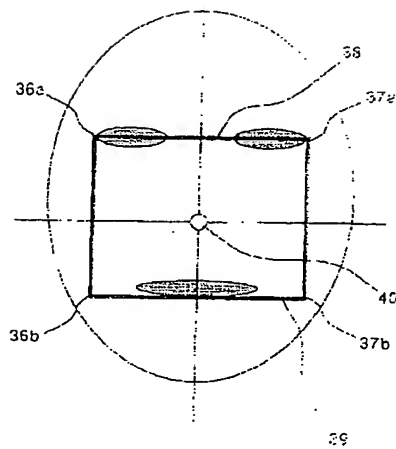
【図16】



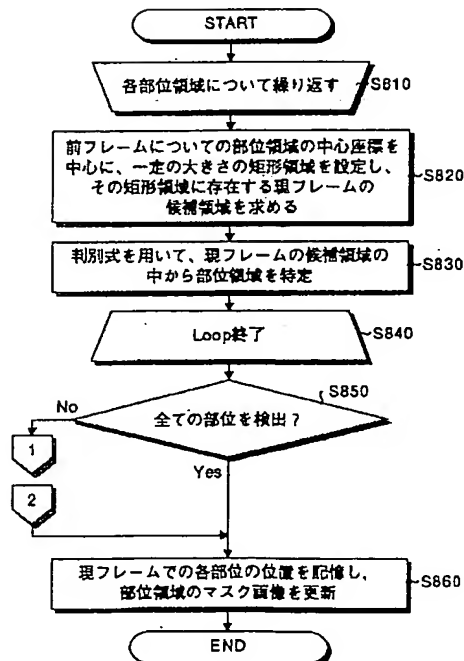
【図17】



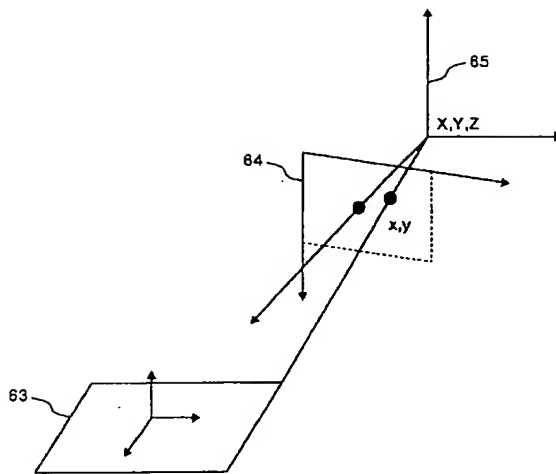
【図18】



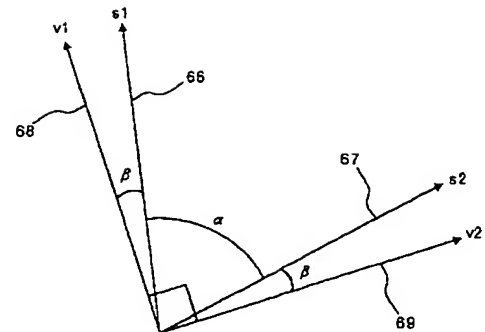
【図20】



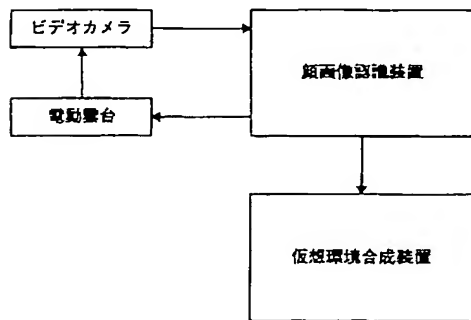
【図28】



【図29】



【図30】



【手続補正書】

【提出日】平成14年10月25日(2002.10.25)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】発明の名称

【補正方法】変更

【補正内容】

【発明の名称】 頭部領域抽出装置およびリアルタイム表情追跡装置

【手続補正2】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

【補正内容】

【特許請求の範囲】

【請求項1】 人物を撮像した映像から人物の頭部領域を抽出する頭部領域抽出装置において、対象人物を撮像した画像の各画素データをR、G、B成分毎に下式 $c1 = \arctan(R / \max(G, B))$

$$c2 = \arctan(G / \max(R, B))$$

$$c3 = \arctan(B / \max(R, G))$$

に従って正規化して正規化データc1、c2、c3を取得する正規化手段と、

正規化データc1、c2、c3を含む各画素データを次式

$$C1 = c2 / c1$$

$$C2 = c3 / c2$$

に従ってC1-C2空間のデータを含む画素データに夫々変換するデータ変換手段と、

変換した画素データのC1データおよびC2データが下式

$$th1 < C1 < th2 \quad th1, th2; \text{肌色抽出パラメータ}$$

$$th3 < C2 < th4 \quad th3, th4; \text{肌色抽出パラメータ}$$

を満足すると、この画素データを肌色画素として判断することにより、撮像した画像から頭部領域を抽出する頭部領域抽出手段と、

【補正対象項目名】0001

【補正方法】変更

【補正内容】

【0001】

【発明の属する技術分野】本発明は、人物を撮像した映像から人物の頭部領域を抽出する頭部領域抽出装置に関する。また、本発明は、本人の顔を送信する代わりにCGキャラクタの映像を相手に送信することによって人物映像を互いに通信するテレビ電話など通信システムに適用され、特にカメラによって撮像された顔の映像から頭部の3次元的な姿勢情報と顔の表情を計測し、この計測結果に基づいてCGキャラクタの動きを制御する代理応答によるリアルタイム表情追跡装置に関するものである。

【手続補正4】

【補正対象書類名】明細書

【補正対象項目名】0009

【補正方法】変更

【補正内容】

【0009】この発明は上記に鑑みてなされたもので、任意の照明条件で撮影された、不特定人物の顔画像から、簡易な演算によって計算能力が低いハードウェアでも実時間で、頭部領域を正確に抽出し得る頭部領域抽出装置を得ることを目的としている。また、この発明は、簡易な演算によって頭部領域を抽出し、頭部の3次元的な動きを計測し、かつ両目および口の開閉状態を計測し、その結果を用いてCGキャラクタの頭部の動きおよび表情を制御するリアルタイム表情追跡装置を得ることを目的としている。

【手続補正5】

【補正対象書類名】明細書

【補正対象項目名】0010

【補正方法】変更

【補正内容】

【0010】

【課題を解決するための手段】上記目的を達成するためこの発明にかかる頭部領域抽出装置は、人物を撮像した映像から人物の頭部領域を抽出する頭部領域抽出装置において、対象人物を撮像した画像の各画素データをR、G、B成分毎に下式

$$c1 = \arctan(R / \max(G, B))$$

$$c2 = \arctan(G / \max(R, B))$$

$$c3 = \arctan(B / \max(R, G))$$

に従って正規化して正規化データc1、c2、c3を取得する正規化手段と、正規化データc1、c2、c3を含む各画素データを次式

$$C1 = c2 / c1$$

$$C2 = c3 / c2$$

に従ってC1-C2空間のデータを含む画素データに夫々変換するデータ変換手段と、変換した画素データのC

1データおよびC2データが下式

$$th1 < C1 < th2 \quad th1, th2; \text{肌色抽出パラメータ}$$

$$th3 < C2 < th4 \quad th3, th4; \text{肌色抽出パラメータ}$$

を満足すると、この画素データを肌色画素として判断することにより、撮像した画像から頭部領域を抽出する頭部領域抽出手段とを備えたことを特徴とする。

【手続補正6】

【補正対象書類名】明細書

【補正対象項目名】0011

【補正方法】変更

【補正内容】

【0011】つぎの発明にかかる頭部領域抽出装置は、上記の発明において、当該対象人物を撮像するときと同じ照明環境下で、対象人物の顔の一部の所定の領域の画像をサンプリングする肌色サンプリング手段と、該肌色サンプリング手段によってサンプリングした所定の領域の画像の各画素データを前記正規化手段を用いて正規化した後、前記データ変換手段を用いてC1-C2空間の画素データに変換し、該変換した前記所定の領域の複数の画素データを用いてC1データについての最大値および最小値とC2データについての最大値および最小値を求め、これらの最大値および最小値で前記肌色抽出パラメータth1、th2、th3およびth4を補正する肌色抽出パラメータ調整手段とをさらに備えることを特徴とする。

【手続補正7】

【補正対象書類名】明細書

【補正対象項目名】0012

【補正方法】変更

【補正内容】

【0012】つぎの発明にかかる頭部領域抽出は、上記の発明において、前記頭部領域抽出手段は、肌色領域抽出結果から最大領域を抽出することにより頭部領域を抽出することを特徴とする。

【手続補正8】

【補正対象書類名】明細書

【補正対象項目名】0013

【補正方法】変更

【補正内容】

【0013】つぎの発明にかかる頭部領域抽出は、上記の発明において、前記頭部領域抽出手段は、頭部領域抽出後の2値画像に膨張収縮処理を加えることを特徴とする。

【手続補正9】

【補正対象書類名】明細書

【補正対象項目名】0014

【補正方法】変更

【補正内容】

【補正方法】削除
【手続補正 16】
【補正対象書類名】明細書
【補正対象項目名】0021
【補正方法】削除
【手続補正 17】
【補正対象書類名】明細書
【補正対象項目名】0022
【補正方法】削除
【手続補正 18】
【補正対象書類名】明細書
【補正対象項目名】0023
【補正方法】削除
【手続補正 19】
【補正対象書類名】明細書
【補正対象項目名】0024
【補正方法】削除
【手続補正 20】
【補正対象書類名】明細書
【補正対象項目名】0025
【補正方法】削除
【手続補正 21】
【補正対象書類名】明細書
【補正対象項目名】0026
【補正方法】削除
【手続補正 22】
【補正対象書類名】明細書
【補正対象項目名】0027
【補正方法】削除
【手続補正 23】
【補正対象書類名】明細書
【補正対象項目名】0028
【補正方法】削除
【手続補正 24】
【補正対象書類名】明細書
【補正対象項目名】0029
【補正方法】削除
【手続補正 25】
【補正対象書類名】明細書
【補正対象項目名】0030
【補正方法】削除
【手続補正 26】
【補正対象書類名】明細書
【補正対象項目名】0031
【補正方法】削除
【手続補正 27】
【補正対象書類名】明細書
【補正対象項目名】0032
【補正方法】変更
【補正内容】

【0032】

【発明の実施の形態】以下に添付図面を参照して、この発明にかかる頭部領域抽出装置およびリアルタイム表情追跡装置の好適な実施の形態を詳細に説明する。このリアルタイム表情追跡装置は、本人の顔を送信する代わりにCGキャラクタの映像を相手に送信することによって人物映像を互いに通信するテレビ電話など通信システムに適用される。

【手続補正 28】

【補正対象書類名】明細書
【補正対象項目名】0136
【補正方法】変更
【補正内容】
【0136】

【発明の効果】以上説明したように、この発明によれば、対象人物を撮像した画像の各画素データをR、G、B成分毎に正規化した後、C1-C2空間のデータを含む画素データに変換し、これら変換した画素データのC1データおよびC2データが所定の肌色抽出パラメータの範囲内に入ると、画素データを肌色画素として判断することにより撮像した画像から頭部領域を抽出するようにしているので、簡易な演算によって肌色成分を極めて正確に抽出することができる。したがって、対象人物の動きに合わせてCGキャラクタの頭部の動きおよび表情を正確に制御することができる。また、使用する照明環境下で対象人物（利用者）の肌色をサンプリングし、このサンプリングデータを用いて肌色抽出のための肌色抽出パラメータを調整するようにしているので、任意の照明環境あるいは利用者毎の個人差に適應して利用者の頭部領域を正確に抽出することができる。

【手続補正 29】

【補正対象書類名】明細書
【補正対象項目名】0137
【補正方法】削除
【手続補正 30】
【補正対象書類名】明細書
【補正対象項目名】0138
【補正方法】削除
【手続補正 31】
【補正対象書類名】明細書
【補正対象項目名】0139
【補正方法】削除
【手続補正 32】
【補正対象書類名】明細書
【補正対象項目名】0140
【補正方法】削除
【手続補正 33】
【補正対象書類名】明細書
【補正対象項目名】0141
【補正方法】削除

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☒ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.